Forecasting Random Walks Under Drift Instability

Hashem Pesaran[1] Andreas Pick[2]

[1]Cambridge University, CIMF

[2]De Nederlandsche Bank

Introduction

- This paper considers the problem of forecasting random walks that are subject to drift and volatility instability.
- Two approaches are considered
  - Averaging over different estimation windows (AveW)
  - Down-weighting observations such that more recent observations carry a larger weight
- We derive asymptotic results for the MSFE for both approaches.
- An empirical application to the monthly growth rate of inflation show that both approaches do improve forecasts.

Literature on Forecasting Under Structural Breaks

- Forecast combinations: Clemen (1989), Stock and Watson (2004), Timmermann (2006)
- Pesaran and Timmermann (2007): Optimal window size
- Aschenmacher-Wesche and Pesaran (2007) and Pesaran, Schuermann and Smith (2007): Applications of averaging over estimation windows (AveW)
- Exponential smoothing: Cox (1961), Gilchrist (1967), Gardner (2006), Branch and Evans (2006)
- Other approaches: Hamilton (1989), Bai and Perron (1998, 2003), Clements and Hendry (2006)

Random walk with a single break

$$y_t = \mu_t + \varepsilon_t, \qquad \varepsilon_t \sim iid(0, \sigma^2)$$

$$\mu_t = \begin{cases} \mu_1, \ \forall\ t \leq T_b, & \text{where } T_b \text{ is the point of break} \\ \mu_2, \ \forall\ t > T_b \end{cases}$$

$$t = 1, 2, \ldots T, \quad T_b < T,$$

Using the entire sample for forecasting, i.e. ignoring the break, we have that

$$\begin{aligned} \hat{y}_{T+h|T} &= \hat{\mu}_{T+h|T} \\ &\to d\mu_2 + (1-d)\mu_1 \end{aligned}$$

where $d = (T - T_b)/T$.

$$\mathrm{E}(y_{T+h} - \hat{y}_{T+h|T}) = (1-d)(\mu_2 - \mu_1)$$

The forecast variance is

$$\mathrm{MSFE}(w) = (1-d)^2(\mu_2 - \mu_1)^2 + \sigma^2/T + \sigma^2$$

AveW forecasts

Observation window:

$$w_{\min} = \frac{T - T_{\min} + 1}{T}$$

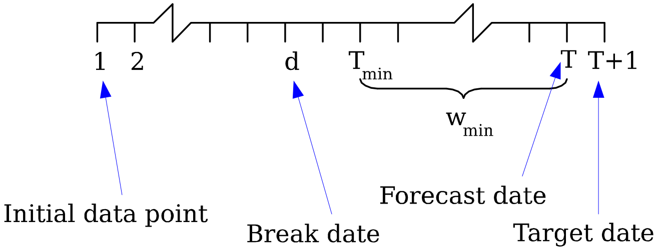
$$w_i = w_{\min} + i/T, \quad i = 0, 1, \ldots, m$$

$$m = T(1 - w_{\min})$$

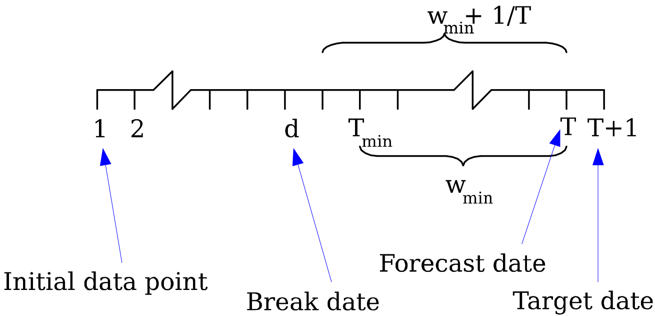
The AveW forecast:

$$\hat{y}_{T+1}(\text{AveW}) = \frac{1}{m+1} \sum_{i=0}^{m} \left(\frac{1}{Tw_i} \sum_{t=T(1-w_i)+1}^{T} y_t \right),$$

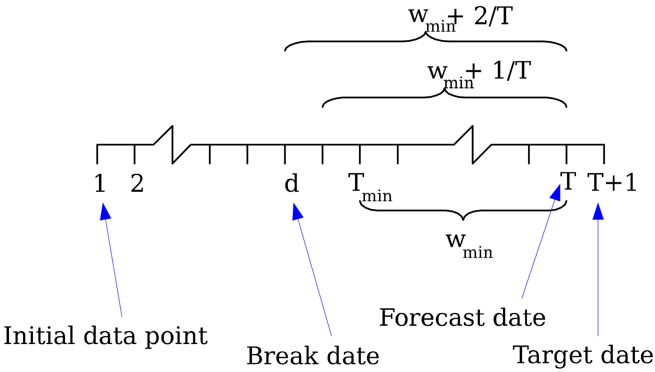
AveW forecasts



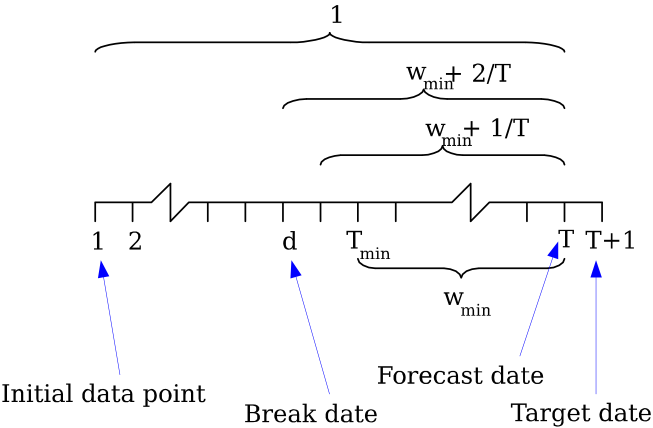
AveW forecasts



AveW forecasts



AveW forecasts



AveW forecasts

For the single window of length wT, the forecast error scaled by σ is

$$\sigma^{-1} e_{T+1}(w) = u_{T+1} + B_{T+1}(w) - \frac{1}{Tw} \sum_{t=T(1-w)+1}^{T} u_t,$$

where the bias is

$$B_{T+1}(w) = \lambda \left(\frac{w-d}{w} \right) \mathrm{I}\left(w-d\right)$$

with $\lambda = (\mu_2 - \mu_1)/\sigma$, and $u_t = \varepsilon_t/\sigma$.

AveW forecasts

For the AveW forecast, the forecast error scaled by σ is

$$\sigma^{-1} e_{T+1}(\text{AveW}) = u_{T+1} + B_{T+1}(\text{AveW}) - \frac{1}{m+1} \sum_{i=0}^{m} \frac{1}{Tw_i} \sum_{t=T(1-w_i)+1}^{T} u_t .$$

where the bias is

$$B_{T+1}(\text{AveW}) = \frac{\lambda}{m+1} \sum_{i=0}^{m} \left(\frac{w_i - d}{w_i} \right) \mathrm{I}\left(w_i - d\right) .$$

The variance of the estimation error

$$\mathrm{V}_{T+1}[\hat{y}_{T+1}(\text{AveW})] = 1 + \left(\frac{1}{T}\right) \left(\frac{1}{m+1}\right)^2 \left[\sum_{i=0}^{m} \frac{1}{w_i} + 2 \sum_{i=0}^{m} \frac{i}{w_i} \right] .$$

AveW forecasts

The difference in MSFE is

$$\begin{aligned} \text{MSFE}(w_a;\lambda,d) - \text{MSFE}(m,w_{\min};\lambda,d) \; &= \\ \lambda^2\left(\frac{w_a-d}{w_a}\right)^2 \text{I}(w_a-d) + \frac{1}{Tw_a} \\ - \left[\frac{\lambda}{m+1}\sum_{i=0}^{m}\frac{w_i-d}{w_i}\text{I}(w_i-d)\right]^2 - \frac{1}{(m+1)^2}\sum_{i=0}^{m}\frac{1+2i}{Tw_i}, \end{aligned}$$

As $T \to \infty$ the sums can be approximated by Riemann integrals, which have closed form solutions.

AveW forecasts

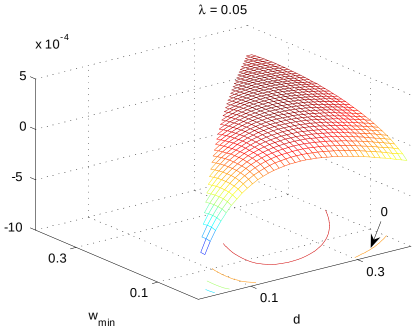
The asymptotic difference in MSFE is

$$
\begin{aligned}
&\mathrm{MSFE}(w=1;\lambda,d)-\mathrm{MSFE}(m,w_{\min};\lambda,d)\\
&=\lambda^2(1-d)^2\,\mathrm{I}(1-d)\\
&\quad-\frac{\lambda^2}{(1-w_{\min})^2}\left[(1-d)+d\ln(d)\right]^2\\
&\quad+O\left(\frac{1}{T}\right)
\end{aligned}
$$

where λ is the size of the break and d the distance to break.

AveW forecasts

Exact difference in MSFE with $T = 100$ and $\lambda = 0.05$



Multiple breaks

Assume

$$\mu_t = \begin{cases} \mu_1, \ \forall\, t \leq T_1 \\ \mu_2, \ \forall\, T_1 > t \leq T_2 \\ \vdots \\ \mu_n, \ \forall\, t > T_n \end{cases}$$

For the single window of size wT the bias is

$$B_{T+1}(w) = \sum_{i=1}^{n} \lambda_i I(w - d_i) \left(\frac{w - d_i}{w} \right),$$

where

$$\begin{aligned} \lambda_i &= \left(\mu_{i+1} - \mu_i\right)/\sigma, i = 1, 2, ..., n \\ n^{-1} \sum_{i=1}^{n} \lambda_i &= \left(\mu_{n+1} - \mu_1\right)/n\sigma. \end{aligned}$$

If $\lambda_1, \lambda_2, ..., \lambda_n$ are distributed independently of $d_1, d_2, ..., d_n$, the bias terms can be approximated for n large as

$$\lim_{n\to\infty} B_F(n) = \mathrm{E}(\lambda_i)(1 - \mathrm{E}(d_i))$$

$$\lim_{n\to\infty} B_{\mathrm{AveW}}(n) = \frac{\mathrm{E}(\lambda_i)}{1 - w_{\min}} \left\{1 - \mathrm{E}(d_i) + \mathrm{E}\left[d_i \ln(d_i)\right]\right\},$$

Therefore, we have

$$\lim_{n\to\infty}\left[B_F^2(n) - B_{\text{AveW}}^2(n)\right] = -\mathrm{E}\left[d_i \ln(d_i)\right]\left\{2 - 2\mathrm{E}(d_i) + \mathrm{E}\left[d_i \ln(d_i)\right]\right\}\mathrm{E}\left[\ln(\lambda_i)\right]$$

and since all parts of the product are non-negative, we have

$$\lim_{n\to\infty}\left[B_F^2(n) - B_{\text{AveW}}^2(n)\right] \geq 0.$$

The strict equality holds only if $E(\lambda_i) = 0$.

For example, if d_i are uniformly distributed over $(0, 1)$, then

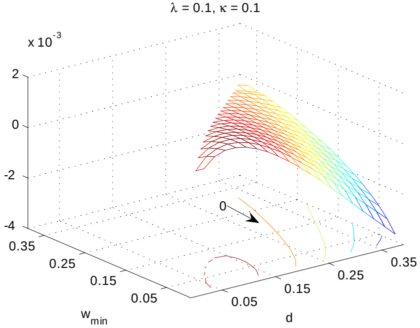
$$\lim_{n\to\infty}\left[B_F^2(n) - B_{\text{AveW}}^2(n)\right] = \frac{3}{16}\left[\mathrm{E}(\lambda_i)\right]^2 \geq 0.$$

Break in drift and volatility

For fixed T, we have that

$$
\begin{aligned}
&\mathrm{MSFE}\left(w;\lambda,d\right)-\mathrm{MSFE}\left(m,s,w_{\min};\lambda,d\right) = \\
= \quad & \lambda^2\left(\frac{w-d}{w}\right)^2 \mathrm{I}(w-d)-\left[\frac{\lambda}{m+1}\sum_{i=0}^{m}\frac{w_i-d}{w_i}\mathrm{I}(w_i-d)\right]^2 \\
& +\frac{w-d}{Tw^2}\mathrm{I}(w-d)\,\kappa^2+\frac{\min(w,d)}{Tw^2} \\
& -\frac{1}{(m+1)^2}\left[\kappa^2\sum_{i=0}^{m}\frac{w_i-d}{Tw_i^2}\mathrm{I}(w_i-d)+\sum_{i=0}^{m}\frac{\min(w_i,d)}{Tw_i^2}\right. \\
& \left.+2\kappa^2\sum_{i=0}^{m-1}\frac{w_i-d}{w_i}\mathrm{I}(w_i-d)\sum_{j=i+1}^{m}\frac{1}{Tw_j}+2\sum_{i=0}^{m-1}\frac{\min(w_i,d)}{w_i}\sum_{j=i+1}^{m}\frac{1}{Tw_j}\right],
\end{aligned}
$$

Exact difference in MSFE for $T = 100$



ExpW: Exponential down-weighting of observations

Consider the Kalman filter

$$\hat{y}_{T+1} = \boldsymbol{\beta}_T' \mathbf{x}_T,$$

$$\boldsymbol{\beta}_T = \boldsymbol{\beta}_{T-1} + \mathbf{G}_T (y_T - \boldsymbol{\beta}_{T-1} \mathbf{x}_{T-1}),$$

If $\mathbf{G}_T = \mathbf{G}$: constant gains updating, which leads to

$$\hat{y}_{T+1}(\text{ExpW}, \gamma) = \hat{y}_{T+1}(\gamma) = \left(\frac{1-\gamma}{1-\gamma^T} \right) \sum_{j=1}^{T} \gamma^{T-j} y_j.$$

Reviews: Gardner (2006) and Branch and Evans (2006)

ExpW: Single break in drift

The forecast bias is

$$\text{Bias}\left[\hat{y}_{T+1}(\text{ExpW},\gamma)\right] = (\mu_2 - \mu_1)\left(\frac{\gamma^{T-T_b+1} - \gamma^T}{1-\gamma^T}\right).$$

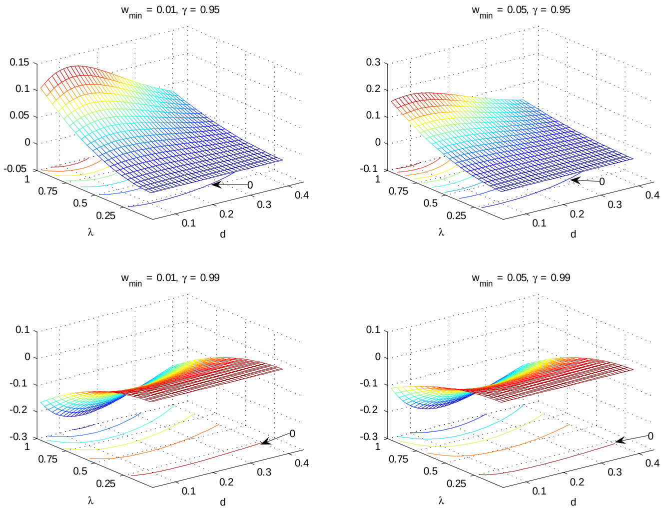
and the forecast error variance is

$$\text{Var}\left[e_{T+1}(\gamma)\right] = \sigma^2\left[1 + \left(\frac{1-\gamma}{1-\gamma^T}\right)^2\left(\frac{1-\gamma^{2T}}{1-\gamma^2}\right)\right].$$

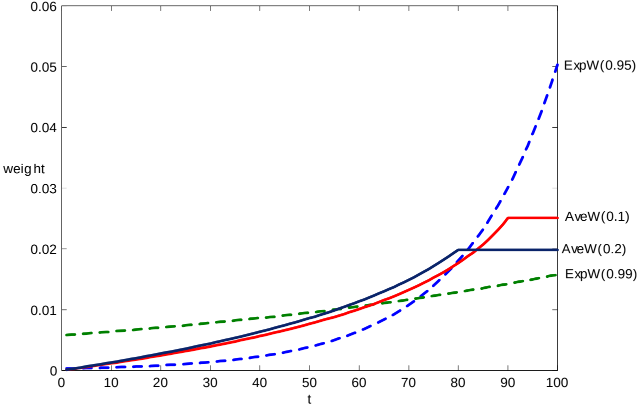
The MSFE is

$$\begin{aligned} \text{MSFE}\left[\hat{y}_{T+1}(\text{ExpW}, \gamma)\right] &= f(\gamma) \quad (1) \\ &= 1 + \lambda^2 \left(\frac{\gamma^{1+T\,d} - \gamma^T}{1-\gamma^T} \right)^2 \\ &+ \left(\frac{1-\gamma}{1-\gamma^T} \right)^2 \left(\frac{1-\gamma^{2T}}{1-\gamma^2} \right), \end{aligned}$$

Comparison of MSFE of AveW and ExpW - $T = 100$



Comparison of weights of observations: AveW and ExpW



Average ExpW over γ

- Weights of ExpW forecast are very sensitive to choice of γ
- For unknown d and λ the optimal choice of γ is unknown
- Robust forecasts may be achieved by averaging ExpW forecasts over a range of γ, say, from 0.95 to 1
- This leads to the averaged ExpW forecast or AveExpW forecast

Application to Forecasting Inflation

- Monthly CPI inflation data
- 21 OECD countries
- 1985M1–2007M10
- 8 year rolling windows, where AveW and ExpW forecasts are calculated for each window
- One month ahead forecasts
- 177 forecasts, with first forecast for 1993M2

Forecast methods

- single window
- ExpW with $w_{\text{min}} = 0.1$
- ExpW with $w_{\text{min}} = 0.2$
- ExpW with $\gamma = 0.95$
- ExpW with $\gamma = 0.98$
- AveExpW with $\gamma \in (0.95, 1)$
- AveExpW with $\gamma \in (0.98, 1)$

- Bias
- RMSFE
- Relative RMSFE

$$\frac{\text{RMSFE(AveW)}}{\text{RMSFE(SW)}},$$

- Diebold-Mariano test for predictive ability for the loss differential

$$l_t(A,B) = e_{tA}^2 - e_{tB}^2,$$

Relative performance in forecasting inflation, $h = 1$

		SW	AveW	ExpW	AveExpW
			0.1	0.95	$(0.95, 1)$
U.S.A.	Bias	0.0253	0.0109	0.0064	0.0144
	RMSFE	0.2261	0.2232	0.2236	0.2237
	SW	1.000	1.6033	0.8253	1.5595
	AveW(0.1)		0.9873	−0.2806	−1.4731
	ExpW(0.95)			0.9890	−0.0732
	AveExpW(0.95, 1)				0.9895
U.K.	Bias	0.0647	0.0284	0.0151	0.0367
	RMSFE	0.3745	0.3675	0.3686	0.3689
	SW	1.000	1.7058	1.0582	1.7877
	AveW(0.1)		0.9813	−0.5537	−1.3588
	ExpW(0.95)			0.9840	−0.1253
	AveExpW(0.95, 1)				0.9849

		SW	AveW	ExpW	AveExpW
			0.1	0.95	(0.95, 1)
Austria	Bias	0.0255	0.0177	0.0131	0.0191
	RMSFE	0.2990	0.2981	0.2991	0.2980
	SW	1.000	0.3070	−0.0347	0.4464
	AveW(0.1)		0.9970	−0.8091	0.1573
	ExpW(0.95)			1.0005	0.6693
	AveExpW(0.95, 1)				0.9967
Belgium	Bias	0.0153	0.0107	0.0078	0.0115
	RMSFE	0.2717	0.2722	0.2741	0.2723
	SW	1.000	−0.2843	−0.9533	−0.4185
	AveW(0.1)		1.0019	−2.0418	−0.1514
	ExpW(0.95)			1.0087	1.5583
	AveExpW(0.95, 1)				1.0021
Canada	Bias	0.0384	0.0175	0.0108	0.0225
	RMSFE	0.3265	0.3198	0.3197	0.3207
	SW	1.000	2.0371	1.4190	2.1992
	AveW(0.1)		0.9794	0.0639	−1.2844
	ExpW(0.95)			0.9791	−0.4692
	AveExpW(0.95, 1)				0.9823

		SW	AveW 0.1	ExpW 0.95	AveExpW (0.95, 1)
Switzerland	Bias	0.0650	0.0383	0.0260	0.0438
	RMSFE	0.3394	0.3326	0.3321	0.3334
	SW	1.000	2.4362	1.9219	2.7404
	AveW(0.1)		0.9801	0.3656	−1.1482
	ExpW(0.95)			0.9785	−0.7586
	AveExpW(0.95, 1)				0.9824
Germany	Bias	0.0275	0.0221	0.0179	0.0228
	RMSFE	0.2018	0.1968	0.1959	0.1967
	SW	1.000	1.3647	1.1899	1.8106
	AveW(0.1)		0.9750	0.5734	0.0978
	ExpW(0.95)			0.9709	−0.3378
	AveExpW(0.95, 1)				0.9746
Denmark	Bias	0.0210	0.0052	0.0010	0.0091
	RMSFE	0.3185	0.3200	0.3230	0.3199
	SW	1.000	−0.6774	−1.3892	−0.7610
	AveW(0.1)		1.0046	−2.4394	0.2780
	ExpW(0.95)			1.0140	2.0922
	AveExpW(0.95, 1)				1.0042

		SW	AveW 0.1	ExpW 0.95	AveExpW (0.95, 1)
Spain	Bias	0.0590	0.0303	0.0201	0.0369
	RMSFE	0.3959	0.3903	0.3916	0.3916
	SW	1.000	2.4339	1.3281	2.3622
	AveW(0.1)		0.9859	−1.1292	−2.3618
	ExpW(0.95)			0.9891	0.0102
	AveExpW(0.95, 1)				0.9891
Finland	Bias	0.0641	0.0274	0.0154	0.0361
	RMSFE	0.3208	0.3076	0.3042	0.3091
	SW	1.000	3.0124	2.6552	3.3707
	AveW(0.1)		0.9587	1.5637	−1.5134
	ExpW(0.95)			0.9482	−1.6822
	AveExpW(0.95, 1)				0.9632
France	Bias	0.0229	0.0108	0.0065	0.0136
	RMSFE	0.2281	0.2253	0.2261	0.2259
	SW	1.000	1.3923	0.7198	1.4167
	AveW(0.1)		0.9879	−0.9074	−1.2095
	ExpW(0.95)			0.9915	0.1805
	AveExpW(0.95, 1)				0.9906

		SW	AveW 0.1	ExpW 0.95	AveExpW (0.95, 1)
Greece	Bias	0.2817	0.1548	0.1082	0.1837
	RMSFE	1.1505	1.1298	1.1323	1.1338
	SW	1.000	1.6476	1.0454	1.7070
	AveW(0.1)		0.9821	−0.4652	−1.4024
	ExpW(0.95)			0.9842	−0.1972
	AveExpW(0.95, 1)				0.9855
Ireland	Bias	−0.0163	−0.0134	−0.0122	−0.0140
	RMSFE	0.3882	0.3871	0.3879	0.3870
	SW	1.000	0.3037	0.0579	0.4220
	AveW(0.1)		0.9973	−0.3814	0.1226
	ExpW(0.95)			0.9992	0.3508
	AveExpW(0.95, 1)				0.9971
Iceland	Bias	0.1418	0.0353	0.0120	0.0631
	RMSFE	0.5517	0.4398	0.4202	0.4583
	SW	1.000	6.2457	5.8937	6.5514
	AveW(0.1)		0.7971	3.7804	−4.7755
	ExpW(0.95)			0.7618	−4.3946
	AveExpW(0.95, 1)				0.8307

		SW	AveW 0.1	ExpW 0.95	AveExpW (0.95, 1)
Italy	Bias	0.0770	0.0428	0.0297	0.0505
	RMSFE	0.1709	0.1519	0.1464	0.1548
	SW	1.000	5.3741	5.0522	5.7647
	AveW(0.1)		0.8888	3.7438	−3.7082
	ExpW(0.95)			0.8566	−3.9352
	AveExpW(0.95, 1)				0.9059
Japan	Bias	0.0453	0.0239	0.0153	0.0286
	RMSFE	0.2370	0.2343	0.2340	0.2344
	SW	1.000	0.7879	0.6989	1.0239
	AveW(0.1)		0.9888	0.2457	−0.0924
	ExpW(0.95)			0.9875	−0.2141
	AveExpW(0.95, 1)				0.9891
Korea	Bias	0.0775	0.0428	0.0272	0.0501
	RMSFE	0.5024	0.5024	0.5045	0.5021
	SW	1.000	0.0046	−0.3552	0.1048
	AveW(0.1)		1.0000	−0.8555	0.2623
	ExpW(0.95)			1.0042	0.8808
	AveExpW(0.95, 1)				0.9993

		SW	AveW 0.1	ExpW 0.95	AveExpW (0.95, 1)
Netherlands	Bias	0.0052	0.0091	0.0064	0.0073
	RMSFE	0.4259	0.4286	0.4318	0.4282
	SW	1.000	−1.0534	−1.6955	−1.1669
	AveW(0.1)		1.0063	−2.6327	0.6103
	ExpW(0.95)			1.0140	2.2872
	AveExpW(0.95, 1)				1.0054
Norway	Bias	0.0588	0.0255	0.0169	0.0339
	RMSFE	0.3188	0.3098	0.3099	0.3111
	SW	1.000	2.7356	1.9518	2.8949
	AveW(0.1)		0.9715	−0.1202	−1.8423
	ExpW(0.95)			0.9721	−0.5869
	AveExpW(0.95, 1)				0.9757
Portugal	Bias	0.1646	0.0875	0.0596	0.1052
	RMSFE	0.4170	0.3725	0.3620	0.3799
	SW	1.000	5.3822	4.8761	5.7041
	AveW(0.1)		0.8932	3.0787	−3.9992
	ExpW(0.95)			0.8681	−3.5852
	AveExpW(0.95, 1)				0.9109

		SW	AveW	ExpW	AveExpW
			0.1	0.95	(0.95, 1)
Sweden	Bias	0.1076	0.0552	0.0358	0.0671
	RMSFE	0.4185	0.3947	0.3895	0.3983
	SW	1.000	3.6392	3.3615	4.0011
	AveW(0.1)		0.9431	2.0109	−2.2729
	ExpW(0.95)			0.9307	−2.3515
	AveExpW(0.95, 1)				0.9517

Summary statistics

- AveW, ExpW, and AveExpW beat the single window in a RMSFE sense in 86% of the forecasts
- For 43% of the forecasts the difference is significant at the 5% level
- The single window forecast is never significantly better
- Overall for this application AveW with $w_{\min} = 0.1$ and ExpW with $\gamma = 0.95$ appear to deliver the best forecasts

Conclusion

- Theoretical results for AveW and ExpW forecasts for the random walk with drift and volatility instability
- AveW and ExpW forecasts always have a lower bias in the presence of structural breaks
- MSFE is also lower if the break exceeds a (very small) value
- Application to CPI inflation shows that this can lead to improved forecasts in practice